Beyond Likes and Loves: Uncovering Complex Sentiments in Bengali Food Reviews on Facebook

Ataf Fazledin Ahamed
Department of Computer Science and Engineering
Bangladesh University of Engineering and Technology
Dhaka, Bangladesh

Sadia Sharmin
Department of Computer Science and Engineering
Bangladesh University of Engineering and Technology
Dhaka, Bangladesh

ABSTRACT

With the rapid growth of internet access, Bangladeshi social media users have been observed to become increasingly active, with Facebook having been established as the primary platform for communication and content sharing. Among various content types, food reviews have been recognized as gaining significant popularity, particularly through video-based posts created by influencers. In this study, the sentiment expressed in Bengali food reviews on Facebook has been investigated through a dual approach: interaction-based sentiment analysis has been performed using Facebook reactions, and sequence classification has been carried out using BanglaBERT, a state-of-the-art Bengali language model. Data have been collected from 905 food review videos and 26,004 associated comments that have been posted between 2020 and 2022. The BanglaBERT model has been fine-tuned on multiple Bengali sentiment datasets, and a prediction accuracy of 83.76% has been achieved, demonstrating its ability to capture nuanced patterns in user opinions. It has been found that the majority of viewers have expressed positive sentiment towards food-related content, although a small subset of comments has reflected contrasting negative opinions. These insights have highlighted the strong engagement value of food reviews in Bangladeshi social media, while the presence of subtle critical perspectives has also been emphasized. A deeper understanding of user perception in this domain has been contributed by the study, and practical implications for content creators, businesses, and researchers in natural language processing have been offered.

Keywords

BanglaBERT, Sentiment Analysis, Bengali Food Reviews, Social Media, Facebook Reactions, Natural Language Processing, Machine Learning

1. INTRODUCTION

The rapid expansion of internet access in the 21st century has fundamentally reshaped the way people have communicated, consumed information, and engaged with digital platforms . As of January 2023, more than 5.16 billion people worldwide have been connected online. In Bangladesh, approximately 38.9% of the population has been reported as internet users[11], with Facebook having been established as the dominant platform for communication, entertainment, and digital interaction. As of 2023, more than 43 million Bangladeshis have been recorded as active users of Facebook, making it the country's most influential social media space [11].

Among the various types of online content, **food-related media**—including reviews, cooking demonstrations, and travel-food videos—has been observed to gain remarkable traction. Such content has been increasingly accessed by large audiences, since it has provided both entertainment and practical value, particularly through mobile devices [6, 5]. The growing presence of **food influencers** in Bangladesh has also been noted, as they have been shown to shape consumer preferences and purchasing behavior. Collaborations with restaurants and businesses have been established, reflecting the commercial importance of food-related content in digital platforms [10].

In this context, the analysis of sentiments expressed by audiences towards food-related content has been considered essential. **Sentiment analysis** has been regarded as a systematic method for capturing public attitudes, ranging from strong appreciation to subtle criticism. For businesses, insights from sentiment analysis have been applied to marketing strategies and customer engagement. For influencers, such insights have been employed to guide content creation. However, despite progress in sentiment analysis for English and other widely spoken languages, research in Bengali has been limited, as the language has been characterized by high complexity and a scarcity of large-scale datasets. Previous studies have mostly been conducted on general comments, hate speech detection, or celebrity-related interactions, while **Bengali food reviews on Facebook have remained underexplored**. Furthermore, the combined use of **Facebook reactions** and advanced models such as **BanglaBERT** for sentiment analysis has not been extensively investigated.

Recent advances (2023–2025) have highlighted the growing attention towards Bengali sentiment analysis. For instance, large-scale resources such as BangDSA and Bangla-Senti have enabled more robust evaluations on social media comments, and fine-tuned transformer models like BanglaBERT have consistently outperformed traditional approaches in benchmark sentiment tasks[7, 9]. Moreover, studies focusing on domain-specific reviews, including food-related datasets, have shown that even classical machine learning models can be competitive when trained on curated corpora[7]. These developments validate the need for nuanced approaches and provide strong motivation for the dual-framework strategy adopted in this paper.

In this study, these gaps have been addressed by adopting a dual approach:

(1) **Interaction-based sentiment analysis** has been conducted using Facebook reactions (e.g., Love, Wow, Angry, Sad) to capture immediate audience responses.
(2) **Sequence classification-based sentiment analysis** has been implemented using BanglaBERT, a state-of-the-art Bengali language model, which has been fine-tuned on multiple sentiment datasets to classify viewer comments.

A dataset of **905 food review videos and 26,004 associated comments**, collected between 2020 and 2022, has been analyzed. It has been found that while the majority of reactions and comments have reflected **positive sentiment**, a small subset has demonstrated contrasting negative perspectives. BanglaBERT has achieved a prediction accuracy of **83.76%**, confirming its effectiveness for real-world sentiment classification.

The key contributions of this work have been summarized as follows:

—A **dual-framework** for sentiment analysis has been proposed, integrating interaction-based features and sequence classification.
—**BanglaBERT** has been fine-tuned on multiple Bengali sentiment datasets and applied to a large corpus of food review comments.
—**Empirical insights** into the sentiment landscape of Bengali food reviews on Facebook have been provided, highlighting positive engagement as well as nuanced negative responses.
—**Practical implications** for food businesses, marketers, and content creators have been demonstrated, supporting the design of effective engagement strategies.

The remainder of this paper has been structured as follows: section literature reviews related work. Section methodology presents the methodology, including data collection and preprocessing. Section result and discussion reports the experimental results and comparative insights. Section conclusion concludes the paper and outlines potential directions for future research.

2. LITERATURE REVIEW

The rise of social media has been accompanied by an explosion of user-generated content, which has been widely recognized as a rich resource for sentiment analysis. Platforms such as Facebook, Twitter, and Instagram have hosted millions of daily interactions, and the analysis of sentiment in these interactions has been considered valuable for businesses, policymakers, and content creators. This section reviews existing research in sentiment analysis with a focus on social media, food-related reviews, Bengali language processing, and interaction-based methods.

2.1 Sentiment Analysis in Social Media

Social media sentiment analysis has been studied extensively, as platforms like Facebook have provided vast amounts of user-generated content that reflect public opinion. Several studies have been conducted to analyze both user reactions (e.g., “Like,” “Love,” “Wow”) and comments, showing how these interactions can be leveraged to measure public perception and engagement.

For example, Khan et al. [12] analyzed Facebook comments from celebrity fan pages to predict emotions such as ‘Happy,’ ‘Angry,’ and ‘Sad.’ Machine learning algorithms such as Random Forest, Support Vector Machines (SVM), Neural Networks, and Naive Bayes have been applied, and it has been demonstrated that these methods can classify emotions in short social media texts. Ishmam and Sharmin [6] investigated hate speech detection in Bengali Facebook posts using Logistic Regression and Gated Recurrent Units (GRUs). Although their work focused on toxic content, the techniques showed broader applicability in sentiment analysis, particularly for noisy and informal social media data. However, classical ML approaches have often struggled with code-mixing, slang, and sparse context, motivating the need for more advanced deep learning and transformer-based models.

Recent works have highlighted these challenges further. Alam et al. [1] introduced **BnSentMix**, a 20,000-sample Bengali-English code-mixed dataset from platforms like Facebook, YouTube, and e-commerce reviews. Baseline experiments with pre-trained transformers achieved only about 69.8% accuracy and 69.1% F1, illustrating the difficulties of code-switching and mixed-script data in sentiment classification. Similarly, Johora et al. [3] presented the **Motamot** dataset of 7,058 politically charged Bangla texts. They compared transformer-based PLMs (BanglaBERT) with large language models (LLMs) such as Gemini 1.5 Pro, finding that few-shot tuned LLMs outperformed PLMs, achieving up to 96.3% accuracy. These studies indicate both the importance of task-specific resources and the emerging role of LLMs in Bengali sentiment analysis.

2.2 Sentiment Analysis in Food Reviews

Sentiment analysis has also been applied to food reviews, where understanding consumer satisfaction is critical. Traditional machine learning algorithms have been used on review platforms such as Zomato, Yelp, and Amazon Foods.

Gite et al. [5] employed Naive Bayes, SVM, and KNN to analyze Zomato and Amazon food reviews, and their results showed the utility of ML for recommendation systems. Islam et al. [10] examined Yelp reviews and confirmed that machine learning methods can effectively classify food-related sentiment. In the Bengali context, Sharif et al. [18] translated Yelp reviews into Bengali and applied a Multinomial Naive Bayes classifier. While this represented an important step toward Bengali food review sentiment analysis, the reliance on translation and classical models limited robustness, particularly in capturing informal expressions typical of user-generated text. Later work extended sentiment analysis to multilingual food reviews, demonstrating the potential of cross-lingual models, which are especially relevant for multilingual countries like Bangladesh. More recently, domain-specific Bengali food review datasets have emerged, where even simple machine learning models achieved competitive performance, underscoring the importance of curated resources for this task [7]. In addition, Rahman and Kabir [17] released the **BDFoodReview** dataset with over 334,000 sentiment-labeled restaurant reviews from Foodpanda Bangladesh in Bangla, English, and Banglish. These contributions show that domain-focused corpora are vital for capturing consumer sentiment in real-world food-related contexts.

2.3 Bengali Language Processing and Sentiment Analysis

Although progress in sentiment analysis for English and other widely spoken languages has been substantial, research in Bengali natural language processing has remained limited due to linguistic complexity and a lack of large-scale datasets. Nevertheless, recent advances have been made.

The introduction of **BanglaBERT** by Bhattacharjee et al. [2] marked a significant advancement, as it outperformed previous models such as mBERT and BanglishBERT on tasks including Named Entity Recognition (NER) and Sentiment Classification (SC). By fine-tuning on large Bengali datasets, BanglaBERT achieved superior performance and established itself as the current benchmark for Bengali sentiment tasks. In addition, Islam et al. [8] introduced the **SentNoB** dataset, designed specifically for noisy Bengali text, which supported model fine-tuning in more realistic conditions. Beyond transformers, earlier studies experimented with Word2Vec embeddings and domain-specific lexicons, which provided contextual signals but generally lagged behind deep transformer models in capturing semantic richness.

Building on this shift, two large-scale resources have expanded the landscape. Islam and Alam [7] introduced **BangDSA**, a dataset of over 200,000 annotated Bangla comments for document-level sentiment analysis, supporting robust evaluation across three-class and 15-class setups. Islam et al. [9] followed with **Bangla-Senti**, a massive collection of 536,930 annotated YouTube comments across diverse domains, where deep learning models, particularly CNNs, outperformed traditional classifiers with an F1 score of 0.86. In addition, Mahmud and Mahmud [14] proposed a hybrid approach combining a lexicon-based polarity score (BSPS) with BanglaBERT, demonstrating improved performance for nine fine-grained sentiment classes over using BanglaBERT alone.

Other research has extended sentiment analysis to specialized settings. The **Motamot** dataset focused on political discourse, showing that large language models (LLMs) such as Gemini 1.5 Pro could surpass traditional PLMs like BanglaBERT with accuracy above 96% in a few-shot setup [3]. Similarly, Alam et al. [1] developed **BnSentMix**, a 20,000-sample code-mixed Bengali-English dataset from social media and e-commerce platforms, where baseline transformers achieved around 70% accuracy—highlighting the challenges of code-switching in sentiment analysis. Collectively, these advances demonstrate that both data scale and model sophistication are essential for handling the unique challenges of Bengali sentiment classification.

2.4 Interaction-Based Sentiment Analysis on Facebook

In addition to text-based approaches, the use of Facebook reactions has been explored as a direct measure of user sentiment. These reactions are valuable since they provide explicit and quantifiable signals.

Freeman et al. [4] demonstrated how reactions can be used to assess sentiment toward scholarly articles by grouping them into positive and negative classes. Pratama [15] applied a similar framework to analyze online learning content during the COVID-19 pandemic. Such approaches indicated that reactions like "Love" or "Wow" can capture strong positive affect, while "Angry" or "Sad" can reflect negative perception. Although promising, these methods have generally been applied in isolation and lacked integration with advanced text classification models, particularly in the Bengali context.

2.5 Research Gaps and Motivation for the Proposed Study

Despite the progress outlined above, several gaps have remained. First, **Bengali food reviews on Facebook have not been extensively studied**, with most prior work focusing on general comments, hate speech, or translated datasets. Second, **interaction-based sentiment analysis** using reactions has rarely been combined with **sequence classification models** such as BanglaBERT to provide a holistic perspective. Third, challenges unique to Bengali—including noisy user input, code-switching with English, and limited annotated corpora—have not been sufficiently addressed in prior research.

The proposed study aims to address these gaps by integrating both approaches: **interaction-based sentiment analysis** using Facebook reactions and **sequence classification-based sentiment analysis** using BanglaBERT. By combining these methods, a more comprehensive understanding of how Bengali-speaking audiences perceive food-related content is achieved. This dual approach ensures that both the linguistic nuances of comments and the immediate emotional responses expressed through reactions are considered.

3. METHODOLOGY

A dual sentiment-analysis approach has been adopted, in which Facebook reaction-based signals have been integrated with sequence classification using BanglaBERT. The dataset has been prepared through staged acquisition, normalization, translation, and linkage to posts; the end-to-end workflow has been shown in Figure 1. Concretely, (1) post data for a fixed time window have been downloaded, (2) ineligible items have been filtered, (3) retained records have been stored as *Main Post Data*, (4) comments for each post have been extracted and formatted, (5) code-mixed text has been translated to Bangla, (6) a further human pass has been performed for quality control and PII masking, (7) a cleaned comment dataset has been created, and (8) comments have been mapped back to their parent posts to yield the *Final Dataset* used in all analyses.

3.1 Dataset Creation

Five high-engagement Bangladeshi food-review influencer pages have been identified using CrowdTangle diagnostics (Fig. 2). Post exports for 2020–2022 have been filtered so that only verified video reviews from official pages have been retained. In the absence of a public comments API, per-post comments have been captured with a browser collection tool, saved as CSV, and linked to the post export via stable `post_id` keys. After de-duplication and integrity checks, **905** videos and **26,004** comments have been retained. This dataset corresponds to the two-stream workflow shown in Fig. 1.

3.1.0.1 Selected pages.. The five public food-review pages are *Rafsan the ChotoBhai*, *Zoltan BD*, *Petuk Couple*, *Khudalagse*, and *MetroMan*. Pages with consistently higher *share of voice* and *interaction rate* have been prioritized so that the corpus has captured the bulk of audience engagement.

3.1.0.2 Engagement snapshot.. Table *3.1.0.2* provides an illustrative snapshot from the left panel of Fig. 2. Values may vary across windows; these figures are representative.

3.1.0.3 Canonical schema.. Post-level fields exposed by CrowdTangle have been summarized in Table 2. Comments have been stored one per row with `comment_id`, `post_id`, timestamp, anonymized author token, and raw text. Personally identifiable information has been masked. To clarify the text-normalization pipeline, illustrative comments and their english forms have been provided in Table 3.

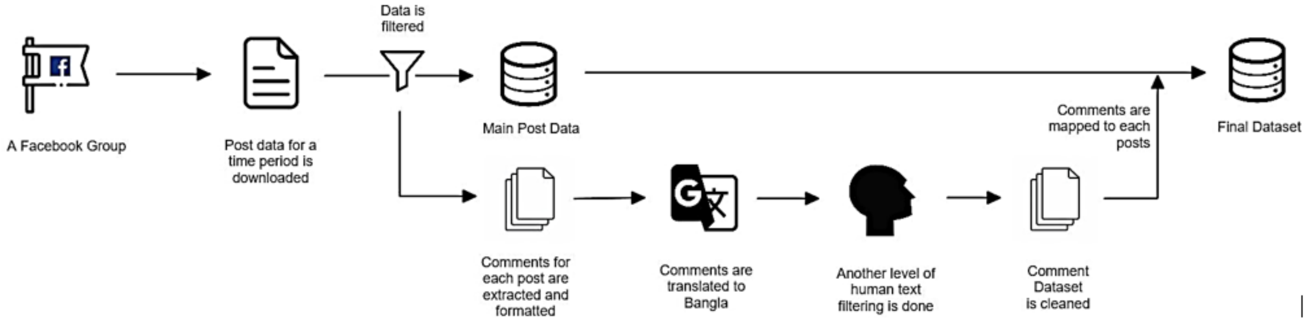

Fig. 1. Workflow of data collection and preprocessing for Bengali food review videos and comments.

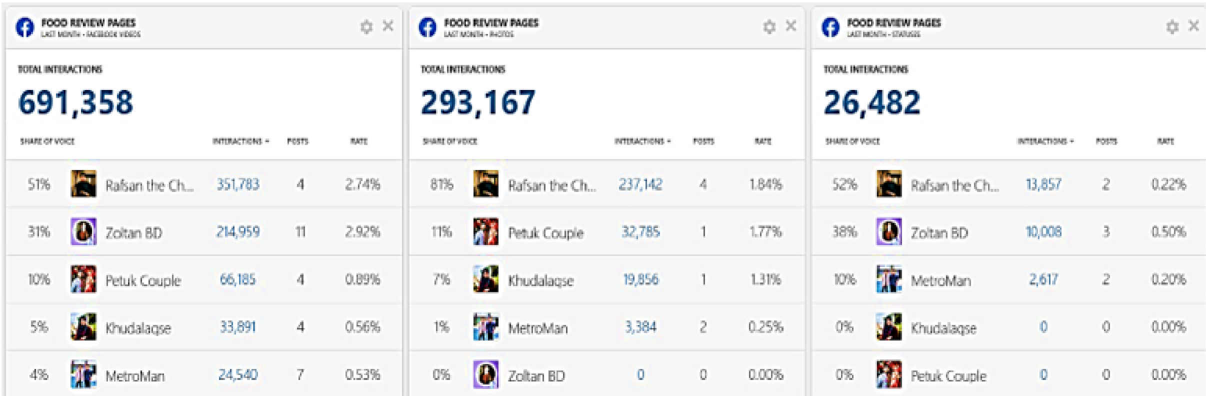

Fig. 2. Page-level interaction diagnostics from CrowdTangle. Each panel summarizes a representative sampling window for the selected food-review pages, showing *Total Interactions*, *Share of Voice*, *Posts*, and *Interaction Rate*. These dashboards have been used to identify high-engagement sources for inclusion.

Table 1. Selected pages and an example engagement snapshot.

Page	**Share of voice (%)**	**Posts in snapshot**	**Notes**
Rafsan the ChotoBhai	51	4	Highest engagement across sampled windows.
Zoltan BD	31	11	High volume with strong interaction share.
Petuk Couple	10	4	Mid-tier engagement; couple-hosted channel.
Khudalagse	5	4	Smaller share; periodic contributions.
MetroMan	4	7	Consistent presence at lower share.

Table 2. Fields available from CrowdTangle for each post

Post Created Date	Post Created Time
Post Type	Total Interactions
Total Comments	Shares
Likes	Love
Wow	Haha
Sad	Angry
Care	Total Views
Video Length	Message
Link	Score

Table 3. Illustrative comments with English translations

Language	**Original comment (masked)**	**English translation**
Bangla	,	The food tasted great, but the service was slow.
English	*The burger was juicy but the bun was stale.*	The burger was juicy but the bun was stale.
Banglish	*price beshi, portion chhoto, abar jabo na.*	The price is high and the portion is small; I will not go again.
Emoji	*Too spicy face savoring food fire but tasty.*	Too spicy but tasty.

3.2 Data Preprocessing

Comments have primarily appeared in **Bangla**, **English**, and **Banglish** (code mixed). A three–stage pipeline has been applied to obtain a single, normalized Bangla stream suitable for modeling. Original and normalized versions have been preserved.

3.2.0.1 Step 1: Banglish to Bangla (Avro phonetic transliteration).. Romanized Bangla strings have been transliterated into standard Bangla script using Avro phonetic rules.

Input (Banglish): *price beshi, portion chhoto, abar jabo na.*
After transliteration: , ,

3.2.0.2 Step 2: English to Bangla translation.. Residual English segments have been translated to Bangla by an automatic MT service. Domain terms have been kept if a faithful translation has not existed.

Input (English): *The burger was juicy but the bun was stale.*
After translation: ,

3.2.0.3 Step 3: Targeted manual correction.. Where automatic output has failed to preserve meaning or tone, minimal human edits have been applied and then propagated by rule.

Auto output: *(colloquial or misspelled)*
Manual correction: *(formal and clearer)*

3.2.0.4 Standard text cleaning.. Light normalization has been performed to retain sentiment while removing noise.

—URLs and trackers have been removed; placeholders have been kept if needed `[URL]`.
—Excess whitespace and repeated punctuation have been collapsed `!!!` $\to$ `!`.
—User tags and emails have been masked `@user` $\to$ `[USER]`.
—Emojis have been replaced by canonical tags during processing $\to$ `[EMO_FOOD][EMO_SPICY]` so that affective cues have been retained. In final exports, tags have been kept or removed depending on the analysis.
—Near duplicates have been removed using character n-gram similarity; the earliest instance has been retained.

3.2.0.5 Worked example (end to end).

Raw comment: *price beshi , portion chhoto, will not recommend!* `http://example.com`
After Step 1: , , will not recommend!
After Step 2: [EMO_FOOD], , !
After cleaning: [EMO_FOOD], , !

The resulting *normalized text* has been used as the input to BanglaBERT, while the *raw text* and the list of applied transformations have been stored to support reproducibility and audit.

3.3 Feature Selection

Two complementary feature sets have been constructed at the post level. The first has captured *interaction signals* derived from grouped reactions; the second has captured *content signals* aggregated from BanglaBERT comment predictions. Empirical motivation has been provided by three figures: reaction volumes (Fig. 3), the correlation matrix (Fig. 5), and the reaction grouping used for polarity (Fig. 4).

A. Features for the Interaction-Based Method. Among the fields received from CrowdTangle, the *Like* count has been discarded for polarity and derived ratios for three reasons: (i) the News Feed ranking algorithm has assigned limited weight to it relative to richer signals [13]; (ii) *Like* has been treated as weak/neutral with low sentiment intensity; and (iii) in this corpus, total *Like* volume has exceeded the sum of all special reactions, which would swamp affective variation if included (Fig. 3). Following prior work [15, 4], reactions have been grouped into **positive** $= \{Love, Care, Wow\}$ and **negative** $= \{Haha, Angry, Sad\}$; this grouping has been visualized in Fig. 4.
Let the grouped totals be

$$n_+ = n_{\text{love}} + n_{\text{care}} + n_{\text{wow}}, \qquad n_- = n_{\text{haha}} + n_{\text{angry}} + n_{\text{sad}}.$$

A post-level polarity score has been computed as

$$Polarity = \frac{n_+ - n_-}{n_+ + n_- + \varepsilon},$$

where a small $\varepsilon > 0$ has prevented division by zero when no qualifying reactions are present. For interpretability, continuous scores have been mapped to ordinal categories in Table 3.3. The correlation matrix in Fig. 5 has shown strong positive coefficients within same-valence pairs (e.g., *Love–Care*, *Haha–Angry*), supporting the grouping.
Derived interaction features. In addition to $Polarity$, several scale-aware and scale-free summaries have been used:

$$s_+ = \frac{n_+}{n_+ + n_- + \varepsilon}, \quad s_- = \frac{n_-}{n_+ + n_- + \varepsilon}, \quad r_{\text{pn}} = \frac{n_+ + \varepsilon}{n_- + \varepsilon},$$

and the reaction-diversity entropy

$$H = - \sum_{r \in \{\text{love,care,wow,haha,angry,sad}\}} p_r \log(p_r), \quad p_r = \frac{n_r}{\sum n_r + \varepsilon}.$$

Robustness. Because *Haha* can be context dependent, a variant in which *Haha* has been treated as neutral (excluded from n_+, n_-) has also been analyzed; conclusions have remained unchanged.

Table 4. Polarity ranges and corresponding categories

Polarity value range	Category
$-1 \leq Polarity < -0.5$	Class 1 (Strong Negative)
$-0.5 \leq Polarity < 0$	Class 2 (Weak Negative)
$0 \leq Polarity < 0.5$	Class 3 (Weak Positive)
$0.5 \leq Polarity \leq 1$	Class 4 (Strong Positive)

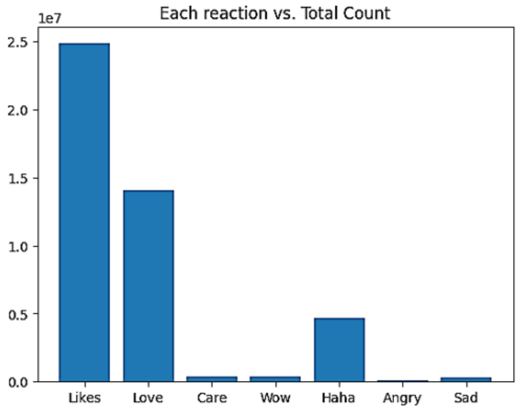

Fig. 3. Reaction volumes by type. *Like* has dominated raw counts, followed by *Love* and a substantial volume of *Haha*.

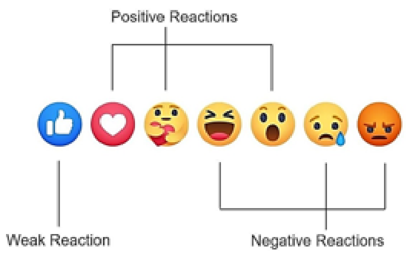

Fig. 4. Grouping of reactions used for polarity. *Like* has been treated as weak/neutral and excluded; {Love, Care, Wow} have formed the positive set; {Haha, Sad, Angry} have formed the negative set.

B. Features for the Sequence Classification-Based Method. Comment sentiment has been inferred with a fine-tuned **BanglaBERT** [2] trained on public Bangla resources [8, 16] and a domain-aligned subset. For each comment i under a post, the model has produced a probability $p_i = \Pr(positive \mid commenti)$ and a hard label $y_i \in \{pos, neg\}$. Let M be the number of comments under a post; the following aggregates have been used:

$$\#pos = \sum_{i=1}^{M} 1[y_i = pos], \qquad \#neg = \sum_{i=1}^{M} 1[y_i = neg],$$

$$Polarity_{cmnt} = \frac{\#pos - \#neg}{\#pos + \#neg + \varepsilon}, \bar{p} = \frac{1}{M}\sum_{i=1}^{M} p_i, c = \frac{1}{M}\sum_{i=1}^{M} |p_i - 0.5|.$$

For categorical summaries, a post-level label has been assigned by the majority of $\{y_i\}$. Directional consistency between comment aggregates and reaction-based polarity has been visible in the correlation matrix (Fig. 5).

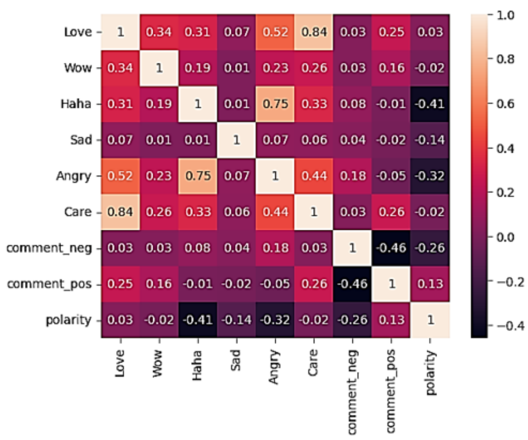

Fig. 5. Correlation matrix among reaction counts, BanglaBERT comment aggregates, and post-level polarity. Same-valence reactions have co-varied; positive comment signals have aligned with polarity while negative comment signals have opposed it.

4. RESULTS AND DISCUSSION

This section has reported quantitative findings from both modalities and has discussed their alignment and implications.

4.1 Model Performance and Benchmarks

Reported benchmarks that have motivated the choice of BanglaBERT within the BERT family have been summarized in Table 5. Study-specific outcomes for the fine-tuned model have been listed in Table 6. Accuracy and Macro-F1 have both been reported to account for label imbalance.

Table 5. Reported performance comparison of BERT-family models on downstream tasks

Models	**Params**	**SC**	**NLI**	**NER**	**BLURB**
mBERT	180M	67.59	75.13	68.97	70.29
BanglishBERT	110M	70.61	80.95	76.28	75.73
BanglaBERT	110M	72.89	82.80	77.78	77.09

Table 6. Performance results of the fine-tuned BanglaBERT used in this study

Metric	**Value**
Validation Accuracy	0.744
Test Accuracy	0.8376
Validation Macro F1	0.699
Test Macro F1	0.8375
Train Loss	0.7038
Validation Loss	0.6400

4.2 Comment-Level Results (BanglaBERT)

The distribution of predicted comment labels has been shown in Figure 6; a predominance of positive comments has been observed. The corresponding histogram of per-video *comment polarity*, aggregating predictions within each video, has been presented in Figure 7. Together, these summaries have indicated that audience responses in the comment stream have tended to be favorable, with a smaller tail of critical views.

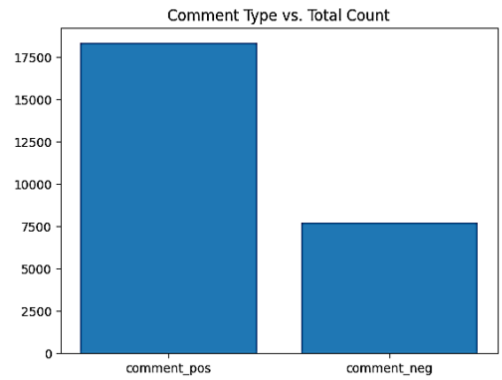

Fig. 6. Distribution of predicted positive and negative comments (BanglaBERT-based). A majority of comments have been predicted as positive.

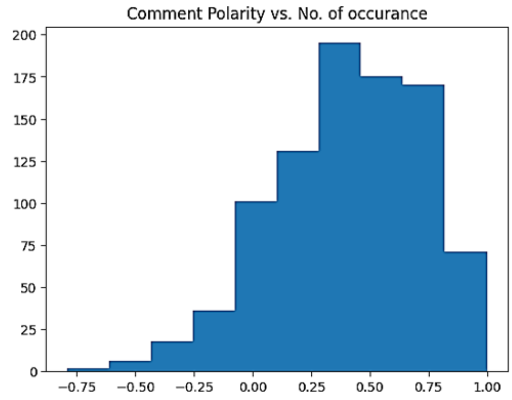

Fig. 7. Histogram of per-video comment polarity derived from BanglaBERT predictions. Most videos have clustered toward positive polarity values.

4.3 Interaction-Based Results (Reactions)

The distribution of videos across interaction-derived polarity categories (Table 4) has been displayed in Figure 8, and the continuous histogram of reaction-derived polarity scores has been shown in Figure 9. A strong skew toward the positive range has been observed. These patterns have persisted when *Like* has been excluded as weak/neutral and under a sensitivity variant in which *Haha* has been treated as neutral.

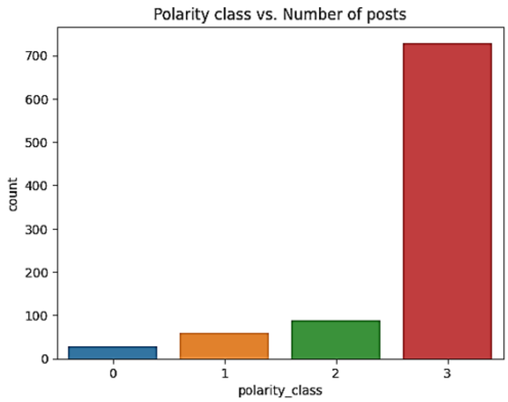

Fig. 8. Distribution of food-review videos by interaction-derived polarity category (see Table 3.3). Positive bins have dominated.

4.4 Comparative Insights

Across both modalities, a predominantly positive audience perception has been observed. The interaction-based approach has provided an aggregate affective signal at the post level, whereas the BanglaBERT-based approach has captured nuanced opinions at the comment level, including minority negative views. The consistency between these views—together with the correlations reported earlier (Fig. 5)—has suggested that reactions and textual sentiment have reflected a common underlying affect toward Bangladeshi food-review videos.

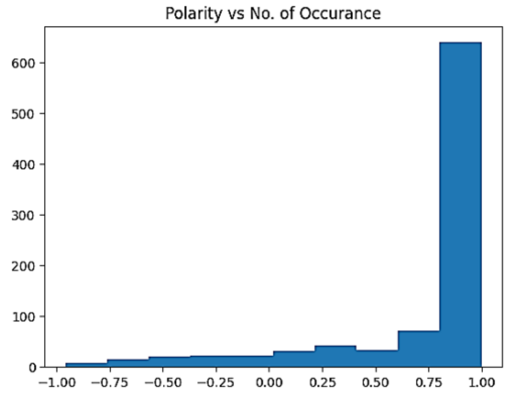

Fig. 9. Histogram of reaction-derived polarity scores across videos. A right shift has indicated overall positive audience affect.

5. CONCLUSION AND FUTURE WORK

A dual strategy for sentiment analysis of Bengali food reviews on Facebook has been investigated, combining an interaction-based approach with sequence classification using BanglaBERT. Across both modalities, a predominantly positive reception has been observed; for example, more than 90% of videos have exhibited positive reaction polarity, and comment-level sentiment has been classified with 83.76% accuracy by the fine-tuned BanglaBERT model. A minority of negative reactions and comments has persisted, indicating areas where critical user perspectives have been expressed.

To broaden coverage and strengthen generalizability, an expansion of the dataset to additional food-review communities and pages has been planned, together with cross-platform collection. The semi-manual data pipeline that has been employed in this study has been slated for further automation and hardening (via the custom browser plugin) to enable scalable, reproducible collection and processing for researchers and practitioners.

The interaction-based sentiment framework inspired by Freeman et al. [15] and extended by Pratama [4] has been identified as a promising direction for richer, low-cost analyses. In future work, multi-class reaction schemes and composite visual summaries (e.g., radar- or profile-based indicators) can be explored to produce more granular audience-affect profiles, while remaining computationally efficient.

Finally, alignment with broader patterns of user engagement on social networks has been noted. Developing sentiment pipelines that are scalable, domain-adaptable, and robust to code-switching, sarcasm, and class imbalance has been prioritized for subsequent research.

6. REFERENCES

[1] Sadia Alam, Md Farhan Ishmam, Navid Hasin Alvee, Md Shahnewaz Siddique, Md Azam Hossain, and Abu Raihan Mostofa Kamal. Bnsentmix: A diverse bengali-english code-mixed dataset for sentiment analysis. *arXiv preprint arXiv:2408.08964*, 2024.

[2] Abhik Bhattacharjee, Tahmid Hasan, Wasi Uddin Ahmad, Kazi Samin, Md Saiful Islam, Anindya Iqbal, M Sohel Rahman, and Rifat Shahriyar. Banglabert: Language model pretraining and benchmarks for low-resource language understanding evaluation in bangla. *arXiv preprint arXiv:2101.00204*, 2021.

[3] Fatema Tuj Johora Faria, Mukaffi Bin Moin, Rabeya Islam Mumu, Md Mahabubul Alam Abir, Abrar Nawar Alfy, and Mohammad Shafiul Alam. Motamot: A dataset for revealing the supremacy of large language models over transformer models in bengali political sentiment analysis. *arXiv preprint arXiv:2407.19528*, 2024.

[4] Cole Freeman, Mrinal Kanti Roy, Michele Fattoruso, and Hamed Alhoori. Shared feelings: Understanding facebook reactions to scholarly articles. In *2019 ACM/IEEE Joint Conference on Digital Libraries (JCDL)*, pages 301–304. IEEE, 2019.

[5] Shilpa Gite, Abhishek Udanshiv, Kartik Jaisinghani, Abhishek Singh, Prafful Chetwani, et al. Sentiment analysis techniques on food reviews using machine learning. In *Proceedings of Second International Conference on Computing, Communications, and Cyber-Security*, pages 511–528. Springer, 2021.

[6] Alvi Md Ishmam and Sadia Sharmin. Hateful speech detection in public facebook pages for the bengali language. In *2019 18th IEEE international conference on machine learning and applications (ICMLA)*, pages 555–560. IEEE, 2019.

[7] Khondoker Ittehadul Islam and Md. Alam. Bangdsa: A large-scale dataset and benchmark for document-level sentiment analysis in bangla. *Findings of the Association for Computational Linguistics: EMNLP*, pages 3265–3271, 2024.

[8] Khondoker Ittehadul Islam, Sudipta Kar, Md Saiful Islam, and Mohammad Ruhul Amin. Sentnob: A dataset for analysing sentiment on noisy bangla texts. In *Findings of the Association for Computational Linguistics: EMNLP 2021*, pages 3265–3271, 2021.

[9] Md. Islam et al. Bangla-senti: A large-scale annotated corpus of 536k youtube comments for sentiment analysis. *arXiv preprint arXiv:2501.01234*.

[10] Nourin Islam, Nasrin Akter, and Abdus Sattar. Sentiment analysis on food review using machine learning approach. In *2021 International Conference on Artificial Intelligence and Smart Systems (ICAIS)*, pages 157–164. IEEE, 2021.

[11] S. Kemp. Digital 2023: Bangladesh — datareportal – global digital insights, 2023. [Online]. Available: `https://datareportal.com/reports/digital-2023-bangladesh`. [Accessed: 20-May-2023].

[12] Md Serajus Salekin Khan, Sanjida Reza Rafa, Amit Kumar Das, et al. Sentiment analysis on bengali facebook comments to predict fan's emotions towards a celebrity. *Journal of Engineering Advancements*, 2(03):118–124, 2021.

[13] Akos Lada, Meihong Wang, and Tak Yan. How does news feed predict what you want to see?, 01 2021. Meta Newsroom.

[14] Hemal Mahmud and Hasan Mahmud. Enhancing sentiment analysis in bengali texts: A hybrid approach using lexicon-based algorithm and pretrained language model bangla-bert. *arXiv preprint arXiv:2411.19584*, 2025.

[15] Ahmad R Pratama. Sentiment analysis of facebook posts through special reactions: the case of learning from home in indonesia amid covid-19. *Jurnal Ilmiah Teknik Elektro Komputer Dan Informatika*, 8(1):83–92, 2022.

[16] Md Atikur Rahman and Emon Kumar Dey. Datasets for aspect-based sentiment analysis in bangla and its baseline evaluation. *Data*, 3(2):15, 2018.

[17] Ehsanur Rahman Rhythm and Md Humaion Kabir Mehedi. Bdfoodreview: A large-scale sentiment-labeled restaurant review dataset from bangladesh. Dataset on Mendeley, 2024. Version 2, 334,119 reviews, mixed Bangla/English/Banglish.

[18] Omar Sharif, Mohammed Moshiul Hoque, and Eftekhar Hossain. Sentiment analysis of bengali texts on online restaurant reviews using multinomial naïve bayes. In *2019 1st international conference on advances in science, engineering and robotics technology (ICASERT)*, pages 1–6. IEEE, 2019.